

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using BW model

Run on: March 29, 2003, 21:10:37 ; Search time 8118.73 Seconds

9201.802 Million cell updates/sec

Title: US-09-988-971-1

Sequence: 1 ccacgcgtccggtcggagc.....aaaaaaaaaaaaaaaaaa 2567

Scoring table: IDENTITY_NUC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Listing first 45 summaries

Database :

1:	gb_ba.*
2:	gb_hcg.*
3:	gb_in.*
4:	gb_om.*
5:	gb_ov.*
6:	gb_pac.*
7:	gb_ph.*
8:	gb_pl.*
9:	gb_pr.*
10:	gb_ro.*
11:	gb_ste.*
12:	gb_sy.*
13:	gb_un.*
14:	gb_vi.*
15:	em_ba.*
16:	em_fun.*
17:	em_hum.*
18:	em_in.*
19:	em_mu.*
20:	em_om.*
21:	em_ov.*
22:	em_or.*
23:	em_pac.*
24:	em_ph.*
25:	em_pl.*
26:	em_ro.*
27:	em_ste.*
28:	em_un.*
29:	em_vi.*
30:	em_hcg_hum.*
31:	em_hcg_inv.*
32:	em_hcg_other.*
33:	em_hcg_mus.*
34:	em_hcg_pin.*
35:	em_hcg_rod.*
36:	em_hcg_nam.*
37:	em_hcg_vrt.*
38:	em_hcg.*
39:	em_hgo_hum.*
40:	em_hgo_mus.*
41:	em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

RESULT 1	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AX452860	2567 bp	DNA	linear	PAT 06-JUL-2002		human.			
Sequence 1 from Patent W00242457.						Homo sapiens			
AX452860						Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AX452860.1	GI:21712520					Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
						Chang, H., Yang, W.P., Wu, Y., Whitney, G.S., Perez-Villar, J.J. and			
						Kramer, S.S.			
						Cloning and expression of human sh2-2: a novel sh2/sh3			

domain-containing human slap homologue having immune cell-specific expression

JOURNAL Patent: WO 0242457-A 1 30-MAY-2002;
Bristol-Myers Squibb Co. (US)

FEATURES Location/Qualifiers
1..2567
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 611 a 741 c 666 g 549 t
ORIGIN

Query Match 100.0%; Score 2567; DB 6; Length 2567;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCCACCGCTCCGCTGGAGCTAGAGCTCCAAAGAACCCAGGCTGTGTCTGTGACAGA 60
1 CCCACCGCTCCGCTGGAGCTAGAGCTCCAAAGAACCCAGGCTGTGTCTGTGACAGA 60
61 GCTCAAAAGGAGCCCTGGGCTTCCCTCCCTGGGCTGTGTGTGGAGGGTTCCCAAGT 120
61 GCTCAAAAGGAGCCCTGGGCTTCCCTCCCTGGGCTGTGTGTGGAGGGTTCCCAAGT 120
121 CCGAATTCCTTAAGAGAGATGGGAGCATGATCCATCCCTGTGTGCAAACTGTGACTG 180
121 CCGAATTCCTTAAGAGAGATGGGAGCATGATCCATCCCTGTGTGCAAACTGTGACTG 180
181 CAGACAGATGCTAGCTACCCAAACCAACAGCTTCCCTGTGAAAGTCTCCCAAGGC 240
181 CAGACAGATGCTAGCTACCCAAACCAACAGCTTCCCTGTGAAAGTCTCCCAAGGC 240
181 CAGACAGATGCTAGCTACCCAAACCAACAGCTTCCCTGTGAAAGTCTCCCAAGGC 240
241 TGAGAGAGTTCTGGGTGTCTTGAAGACCAAGAGACATGGGAGACTTCCAAAGAGGCCCA 300
241 TGAGAGAGTTCTGGGTGTCTTGAAGACCAAGAGACATGGGAGACTTCCAAAGAGGCCCA 300
301 AAGCCCTAACCTGTCCAGCCAGAGATGGGTCTCAGAGAGCTGTCTTCCCAAGCCTTTG 360
301 AAGCCCTAACCTGTCCAGCCAGAGATGGGTCTCAGAGAGCTGTCTTCCCAAGCCTTTG 360
361 ATGCAAAACCAATTTCCCTGATGATGATGCTTCTGAGGCTCTGTGAGAGAAACATGGA 420
361 ATGCAAAACCAATTTCCCTGATGATGATGCTTCTGAGGCTCTGTGAGAGAAACATGGA 420
421 AGTCTGCCAGCAGAAAGAAATCTTCCCAAGCCCAAGCTTGAAGTTCTGTGTCCAAAGC 480
421 AGTCTGCCAGCAGAAAGAAATCTTCCCAAGCCCAAGCTTGAAGTTCTGTGTCCAAAGC 480
481 CAGGAGACTGTGACATGAG 540
481 CAGGAGACTGTGACATGAG 540
541 TTCCCGGAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGAGCCATTGACATCGTGC 600
541 TTCCCGGAGGTGGCCCGGCGAGCTGTGCTGAGACTCGGGAGAGCCATTGACATCGTGC 600
601 TCTGAGATGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
601 TCTGAGATGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
661 CCCAGGCTCCAGTGGCCAAAGTCTCCATGGGTGGCTGTATATAGAGGCTGTGAGAGAG 720
661 CCCAGGCTCCAGTGGCCAAAGTCTCCATGGGTGGCTGTATATAGAGGCTGTGAGAGAG 720
721 AAAGCAGAGAACTGTGTGTTTACCTGTGGAACCTGTGAGAGAGAGAGAGAGAGAGAG 780
721 AAAGCAGAGAACTGTGTGTTTACCTGTGGAACCTGTGAGAGAGAGAGAGAGAGAGAG 780
781 AGCCAGACAGAGAGAGCTCTTACTCTGTCAATGCTGAGAGAGAGAGAGAGAGAGAG 840
781 AGCCAGACAGAGAGAGCTCTTACTCTGTCAATGCTGAGAGAGAGAGAGAGAGAGAG 840
841 GACCGATCAGACATCAGAGATCAGTGTGCAATGAGCTGGGCTGACATCTCACCG 900
841 GACCGATCAGACATCAGAGATCAGTGTGCAATGAGCTGGGCTGACATCTCACCG 900

841 GACCGATCAGACATCAGAGATCAGTGTGCAATGAGCTGGGCTGACATCTCACCG 900
901 CGGCTACCTTCCCTCTCACTCCAGAGCCCTGTGTGAGACATTAATCTGTAGCTGGGATGAC 960
901 CGGCTACCTTCCCTCTCACTCCAGAGCCCTGTGTGAGACATTAATCTGTAGCTGGGATGAC 960
961 ATGTGTCTCTCACTCAAG 1020
961 ATGTGTCTCTCACTCAAG 1020
1021 GATATACCTCTCACTCAAG 1080
1021 GATATACCTCTCACTCAAG 1080
1021 GATATACCTCTCACTCAAG 1080
1081 TCCCTCTCTCTCTCAAG 1140
1081 TCCCTCTCTCTCTCAAG 1140
1141 GAGTCTCTCTCTCTCAAG 1200
1141 GAGTCTCTCTCTCTCAAG 1200
1201 GGCCTCAAG 1260
1201 GGCCTCAAG 1260
1261 CTTGGGAG 1320
1261 CTTGGGAG 1320
1321 ATCTAG 1380
1321 ATCTAG 1380
1381 TCTTCTCAAG 1440
1381 TCTTCTCAAG 1440
1441 AGGAGCAG 1500
1441 AGGAGCAG 1500
1501 AGTTGGGAG 1560
1501 AGTTGGGAG 1560
1561 ACCATTAAACAG 1620
1561 ACCATTAAACAG 1620
1621 TTTTCAAGAAATCTCAAG 1680
1621 TTTTCAAGAAATCTCAAG 1680
1681 GGCCTTAAG 1740
1681 GGCCTTAAG 1740
1741 TCCAGGCTCTATGAG 1800
1741 TCCAGGCTCTATGAG 1800
1801 ATGTCCCTCACTTGAAG 1860
1801 ATGTCCCTCACTTGAAG 1860
1861 GGCCTGTGGAG 1920
1861 GGCCTGTGGAG 1920
1921 CTGATTTCAAG 1980
1921 CTGATTTCAAG 1980

QY 1981 CACCATGCGCTGCTAATTTTGTGATTTTGTAGTACAGGGGGTTTACACATTTGGCC 2040
 DB 1981 CACCATGCGCTGCTAATTTTGTGATTTTGTAGTACAGGGGGTTTACACATTTGGCC 2040
 QY 2041 AGGCTGTGTGAACTCTGACCTGAGGTGATCCACCACTTTGGCTCCCAAGTGTG 2100
 DB 2041 AGGCTGTGTGAACTCTGACCTGAGGTGATCCACCACTTTGGCTCCCAAGTGTG 2100
 QY 2101 GGATTTACAGGTGAGCCAGCCGACCTGACCTGATCTCTATTTTCTTTTGG 2160
 DB 2101 GGATTTACAGGTGAGCCAGCCGACCTGACCTGATCTCTATTTTCTTTTGG 2160
 QY 2161 GCTTACATTCCTTACGACATCTGGCTTGCATTTGTGGCCGAATTAATAATACACT 2220
 DB 2161 GCTTACATTCCTTACGACATCTGGCTTGCATTTGTGGCCGAATTAATAATACACT 2220
 QY 2221 CTTAAGCTACACACTGAGTGAAGCCAGGACCTTCACTGTGGGAGGGGCAATGAA 2280
 DB 2221 CTTAAGCTACACACTGAGTGAAGCCAGGACCTTCACTGTGGGAGGGGCAATGAA 2280
 QY 2281 GGTTGTAAGCCCTCTCTCAATGCCAAGACGAGACCAAGCTTACACCAATCCAGC 2340
 DB 2281 GGTTGTAAGCCCTCTCTCAATGCCAAGACGAGACCAAGCTTACACCAATCCAGC 2340
 QY 2341 CCTGATTTCCCTGCTGCTCCATTAACGAAAGAGTCTGCTGATTCCTGTAAGGATC 2400
 DB 2341 CCTGATTTCCCTGCTGCTCCATTAACGAAAGAGTCTGCTGATTCCTGTAAGGATC 2400
 QY 2401 AGGAGAGAGAAAGAGGATGGGGTGGAGGACCCCTCACTGCTCTCTGTTTC 2460
 DB 2401 AGGAGAGAGAAAGAGGATGGGGTGGAGGACCCCTCACTGCTCTCTGTTTC 2460
 QY 2461 CCAAGCTACAGGTGGGGTGGGAAAGCTTTATCGAGTATCAACAGTTTCATTTAA 2520
 DB 2461 CCAAGCTACAGGTGGGGTGGGAAAGCTTTATCGAGTATCAACAGTTTCATTTAA 2520
 QY 2521 AGATTGATTTATTCAGATGTGAAAAAAGAAAAAAGAAAAA 2567
 DB 2521 AGATTGATTTATTCAGATGTGAAAAAAGAAAAAAGAAAAA 2567

RESULT 2
 LOCUS AK025645 2415 bp mRNA linear PRI 29-SEP-2000
 DEFINITION Homo sapiens cDNA: FLJ21992 fls, clone HEP06554.
 ACCESSION AK025645
 VERSION AK025645.1 GI:10438227
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens hepatoma cell_line: HepG2 cDNA to mRNA, clone lib:HEP
 clone:HEP06554.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE Kawaaba, A., Hiki, T., Kobayashi, N., Inagaki, H., Ikemura, Y.,
 Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
 Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
 NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2415)
 AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T.,
 Shibahara, T., Tanaka, T. and Nakamura, Y.
 TITLE Direct Submission 2/4
 JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
 Japan (E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
 Fax: 81-3-5449-5416)

COMMENT
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction, 5'- & 3'-end one pass sequencing: Department of

Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).

FEATURES
 source Location/Qualifiers
 1. 2415
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEP06554"
 /cell_line="HepG2"
 /cell_type="hepatoma"
 /clone_id="HEP"
 /note="Cloning vector pHEP18FLJ3"
 61. 846
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB15201.1"
 /db_xref="GI:10438228"

CDS

BASE COUNT 599 a 662 c 618 g 536 t
 ORIGIN
 Query Match 84.2%; Score 2160.2; DB 9; Length 2415;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 2179; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 372 ATTTCCTCATATGCTCTCTGAGTGTCTGAGGAAACATGGAAAGCTGCCAG 431
 DB 18 ATTTCCTCATATGCTCTCTGAGTGTCTGAGGAAACATGGAAAGCTGCCAG 77
 QY 432 CAGAAAGAAATCTTCCCAAGCCCAAGCTTGAATCTCTGTCAGAGCCAGCTGT 491
 DB 78 CAGAAAGAAATCTTCCCAAGCCCAAGCTTGAATCTCTGTCAGAGCCAGCTGT 137
 QY 492 GACCATGAGAGCAG 551
 DB 138 GACCATGAG 197
 QY 552 TGGCCCGCCGAGCTGTGCTGAGATCTGGGAGCCATTGACATGTCTTGAAGATGG 611
 DB 199 TGGCCCGCCGAGCTGTGCTGAGATCTGGGAGCCATTGACATGTCTTGAAGATGG 257
 QY 612 AGACTGTGAGACGAGTGTCTGAAAGTCTCAGAGAGAGATATATCCCAAGCTCCA 671
 DB 258 AGACTGTGAGACGAGTGTCTGAAAGTCTCAGAGAGAGATATATCCCAAGCTCCA 317
 QY 672 CGTGGCCAAAGTCTCCATGAGGTGCTGTATGAGGCTTGACAGAGAGAGAGAGAG 731
 DB 318 CGTGGCCAAAGTCTCCATGAGGTGCTGTATGAGGCTTGACAGAGAGAGAGAGAG 377
 QY 732 ACTGCTGTATTACTCTGGGAAACCTTGAAGGAGGCTTCTCATCCGGAGAGCAGACAG 791
 DB 378 ACTGCTGTATTACTCTGGGAAACCTTGAAGGAGGCTTCTCATCCGGAGAGCAGACAG 437
 QY 792 GAGAGGCTCTTACTCTGTGATGCTGAGCTCAGCCGCTTCATCTGGAGCCGATAG 851
 DB 438 GAGAGGCTCTTACTCTGTGATGCTGAGCTCAGCCGCTTCATCTGGAGCCGATAG 497
 QY 852 ACACTACAGATCACTGCTTGAATGAGGCTGAGTCAATCTCAGCCGCTCACTT 911
 DB 498 ACACTACAGATCACTGCTTGAATGAGGCTGAGTCAATCTCAGCCGCTCACTT 557
 QY 912 CCCCTCACTCAGAGCCCTGAGTGAACATTAATGAGTGGCGGATGACATCTGTGCT 971
 DB 558 CCCCTCACTCAGAGCCCTGAGTGAACATTAATGAGTGGCGGATGACATCTGTGCT 617
 QY 972 ACTCAAGAGCCCTGTGCTGAGAGGCTGAGCCGCTCCCTGAGAGAGATATACCTT 1031
 DB 618 ACTCAAGAGCCCTGTGCTGAGAGGCTGAGCCGCTCCCTGAGAGAGATATACCTT 677

Chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

This sequence is the entire insert of clone RP5-977B1. The true left end of clone RP3-460B8 is at 6345 in this sequence. The true right end of clone CTD-2182L9 is at 6218 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP5-977B1 is from the library RPT-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

source

Location/Qualifiers

1. .145068
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP5-977B1"
/clone_1ib="RPT-5"
568. .875
/note="AluSc repeat: matches 1. .311 of consensus"
1024. .1580
/note="match: GSS: Em:AQ620498"
1047. .1629
/note="match: GSS: Em:AQ779162"
1631. .1923
/note="AluSc repeat: matches 1. .294 of consensus"
2776. .2842
/note="AluSc repeat: matches 2357. .2423 of consensus"
3891. .3949
/note="MIR repeat: matches 34. .90 of consensus"
4458. .4611
/note="MIR repeat: matches 67. .231 of consensus"
5133. .5549
/note="L1MD2 repeat: matches 5924. .6340 of consensus"
5550. .5680
/note="L1MD2 repeat: matches 1. .132 of consensus"
5703. .5890
/note="AluSc repeat: matches 120. .308 of consensus"
5923. .6012
/note="L1MC/D repeat: matches 5432. .5528 of consensus"
6588. .6889
/note="AluSc repeat: matches 1. .303 of consensus"
join(7423. .7784,9960. .10039,10306. .10394,10917. .11329,34968. .35059,36569. .36724,37531. .39351)
/gene="KIAA0964"
/product="dJ977B1.6.2 (KIAA0964 ortholog of rat PSD-95/SAP90-associated protein 4)"
/note="match: CDNAS: Em:AB023181. Em:AF009204. Em:AB003594. Em:U67988. Em:U67137. Em:U67138. Em:U67987. Em:U67139. Em:AF131778. Match. ESTs: Em:AI81867. Em:AI20016. Em:AA496750. Em:AA647532. Em:AI747584. Em:AA931146. Em:AI887761. Em:AA154844. Em:AA278057. Em:AI202736. Em:AI97978. Em:AA191427. Em:AA611645. Em:AA760107. Em:AA599160. Em:AA446940. Em:AA190916. Em:T07912. Em:AA558202. Em:AI748854. Em:AI316293. Em:AI840413. Em:AI347117. Em:AA571014. Em:AA870776. Em:AA144662. Em:AA126125. Em:AA496805. Em:W62099

gene

mRNA

mRNA

CDS

CDS

Em:AA443364 Em:AA874407 Em:AI986170 Em:AI840039
Em:AA943725 Em:AA895674 Em:AI493631 Em:AA760121
Em:AA050197 Em:R84650 Em:T35758 Em:R06650 Em:AA433898
Em:R54251 Em:AI918247 Em:H05433 Em:W42663 Em:T75108
Em:AA823397 Em:R13930 Em:AI949469 Em:AA861813 Em:AA602005
Em:AI201639 Em:W45187 Em:Z45916 Em:R12331 Em:AI936733
Em:R12936 Em:R36776 Em:HI2916 Em:R13255 Em:AI983945
Em:AA126063 Em:T57999 Em:AI916248 Em:T75144 Em:W07614
Em:AI972603 Em:AA434575 Em:AI500286 Em:AI201649
Em:AI343454 Em:AI128723 Em:AI179528 Em:AA031035 Em:T08824
Em:N52262 Em:R51856 Em:R96064 Em:Z42146 Em:T49317
Em:R96101 Em:HI3361 Em:AA064703 Em:AA295898 Em:AI049119
Em:AA222576 Em:D28675 Em:AI652021 Em:AA082413 Em:H05501
Em:NA4297 Em:AA046329 Em:W45937 Em:AA577212 Em:AA610855
Em:AI017860 Em:W8510 Em:T08842 Em:AI057148 Em:W28464
/evidence=not_experimental
join(<7423. .7784,9960. .10039,10917. .11329,34968. .35059,36569. .36724,37531. .39351)
/gene="KIAA0964"
/note="isoform 2
continues in dJ996C2 (AL445705) and BK2182L9 (AL390374)
match: proteins: Tr:Q9Y2H0"
/codon_start=3
/evidence=not_experimental
/product="dJ977B1.6.2 (KIAA0964 ortholog of rat PSD-95/SAP90-associated protein 4)"
/protein_id="CAC16960.1"
/db_xref="GI:1132333"
/db_xref="SPTRMBL:O9H138"
/translation="SSCIAYAKKPPVPVPPRTSKPFIYVQSTESADDTYLDSD
HKSRYTSGISNNSDLSSTPSPVTRGQVPAPEAPPPKAAAKSKSGTLTS
ESHPPAKRKLRSIGIQRTRNSHLSIDAGRPAPKIAQIKNLVYGNDDPA
TSDTDQNDNSCKSSERSLPDCTPHNPSISIDAGRPAPKIAQIKNLVYGNDDPA
LBSASLPDPDWLETSSSPAPAPGACRDYWLKLAQETERLEGCCOMDET
KENNISSEVLGKLVSAVGSQAOLMSQFOFQRCGLQNLNPNAPPTADLDGFPV
LQSLIEDISMKFDELVHLKANSWOLVETPEKREKKPPVPVKKPKAPASRDKA
SDASDKORAEKRLIAAKRAASVRQNSATESADSIETVPEAOTRL"
join(<7423. .7784,10306. .10394,10917. .11329,34968. .35059,36569. .36724,37531. .39351)
/gene="KIAA0964"
/note="isoform 1
continues in dJ996C2 (AL445705) and BK2182L9 (AL390374)
match: proteins: Tr:O14488 Tr:O14664 Tr:O14489 Tr:P97839
Tr:O14490 Tr:P97836 Tr:P97837 Tr:O54773 Tr:P97841
Tr:P97835"
/codon_start=3
/evidence=not_experimental
/product="dJ977B1.6.1 (KIAA0964 ortholog of rat PSD-95/SAP90-associated protein 4)"
/protein_id="CAB5370.1"
/db_xref="GI:6983730"
/db_xref="SPTRMBL:O9H137"
/translation="SSCIAYAKKPPVPVPPRTSKPFIYVQSTESADDTYLDSD
HKSRYTSGISNNSDLSSTPSPVTRGQVPAPEAPPPKAAAKSKSGTLTS
ESHPPAKRKLRSIGIQRTRNSHLSIDAGRPAPKIAQIKNLVYGNDDPA
LBSASLPDPDWLETSSSPAPAPGACRDYWLKLAQETERLEGCCOMDET
KENNISSEVLGKLVSAVGSQAOLMSQFOFQRCGLQNLNPNAPPTADLDGFPV
LQSLIEDISMKFDELVHLKANSWOLVETPEKREKKPPVPVKKPKAPASRDKA
SDASDKORAEKRLIAAKRAASVRQNSATESADSIETVPEAOTRL"
10065. .10112
/note="6 copies 8 mer cgcctgc 79% conserved"
11537. .11753
/note="MIR repeat: matches 11. .236 of consensus"
11767. .11782
/note="L2 repeat: matches 2692. .2708 of consensus"
11783. .12083
/note="AluSc repeat: matches 1. .300 of consensus"
12084. .12121
/note="L2 repeat: matches 2661. .2692 of consensus"
12631. .12714
/note="HAI1 repeat: matches 1589. .1665 of consensus"
13077. .13367

```

repeat_region /note="AluJo repeat: matches 3. .290 of consensus"
13440. /note="AluX repeat: matches 1. .304 of consensus"
repeat_region /note="AluX repeat: matches 1. .14197
13332. /note="L1M8I repeat: matches 5915. .6173 of consensus"
repeat_region /note="L1M8I repeat: matches 14198. .14512
14198. /note="AluY repeat: matches 1. .307 of consensus"
repeat_region /note="L1M8I repeat: matches 14513. .14756
14513. /note="L1M8I repeat: matches 5650. .5915 of consensus"
repeat_region /note="L1M8I repeat: matches 14789. .15705
14789. /note="L1M8A repeat: matches 5162. .6185 of consensus"
repeat_region /note="L1M8A repeat: matches 15706. .16004
15706. /note="AluSg repeat: matches 1. .298 of consensus"
repeat_region /note="L1M8A repeat: matches 16005. .16060
16005. /note="L1M8A repeat: matches 5111. .5162 of consensus"

```

Query Match	57.0%	Score 1462.6;	DB 9;	Length 145068;
Best Local Similarity	99.3%	Pred. No. 0;		
Matches 1479;	Conservative	0;	Mismatches 9;	Indels 1;
			Gaps	1;

Oy	1076	ACAGCCTCCCTCTGTTTTCGAAAGCTGCACAGGGAGGAAGTCCTTCTCAGTGAAGGGTC	1135
Db	124708	ACAGCTCCCTCTCTGTTTTCGAAAGCTGCACAGGGAGGAAGTCCTTCTCAGTGAAGGGTC	1246459
Oy	1136	TCCGGAGATCCCTCAGCTTCTACATCAGCTGAATACAGAGCTGCTCTTTTGATGATG	1195
Db	124648	TCCGGAGATCCCTCAGCTTCTACATCAGCTGAATACAGAGCTGCTCTTTTGATGATG	1245899
Oy	1196	CCTAGGCCCAAGAGAGAGGGCAAAAGGGAAACCAAGCTGCACACTAGAAACCCCAATTC	1255
Db	124588	CCTAGGCCCAAGAGAGAGGGCAAAAGGGAAACCAAGCTGCACACTAGAAACCCCAATTC	1245299
Oy	1256	AGCTTCCTGGGCACCCCAAGGCAGAGCTGTGCACTCAGGGAGGGAGGTGGACACAGA	1315
Db	124528	AGCTTCCTGGGCACCCCAAGGCAGAGCTGTGCACTCAGGGAGGGAGGTGGACACAGA	1244695
Oy	1316	GGGCGATCTAAGAGGTCCCACTGTGAACCTTGTCTTCTCTTTAGCCCTTAGAGTAC	1375
Db	124468	GGGCGATCTAAGAGGTCCCACTGTGAACCTTGTCTTCTCTTTAGCCCTTAGAGTAC	1244090
Oy	1376	CTACTTCCTTCCAGTGCCATGATCCCACTTGCACTCTAGTGGAGTGCAGAGAAAGTG	1435
Db	124408	CTACTTCCTTCCAGTGCCATGATCCCACTTGCACTCTAGTGGAGTGCAGAGAAAGTG	1243490
Oy	1436	GGACACAGGGCCAGGGTTCAAAAAGAAATAGAGCTCTGGGGGGTCTGACTAGTTAGT	1495
Db	124348	GGACACAGGGCCAGGGTTCAAAAAGAAATAGAGCTCTGGGGGGTCTGACTAGTTAGT	1242890
Oy	1496	TCTTGAGTTGGGGTTTCCAGTACCATCTGGAATGCCCTGCTTTGAGCCCCATTTCACA	1555
Db	124288	TCTTGAGTTGGGGTTTCCAGTACCATCTGGAATGCCCTGCTTTGAGCCCCATTTCACA	1242290
Oy	1556	TCCGCACATTAAACAGAGGCCCAACCAAGTATGAAACACCCCTAGAGTCAACAGAA	1615
Db	124228	TCCGCACATTAAACAGAGGCCCAACCAAGTATGAAACACCCCTAGAGTCAACAGAA	1241690
Oy	1616	AGTCATTTTCAAGAAATCTACAAATCTGTTGAGACCAACCATCACTCAGAAAGGTAG	1675
Db	124168	AGTCATTTTCAAGAAATCTACAAATCTGTTGAGACCAACCATCACTCAGAAAGGTAG	1241090
Oy	1676	ACTGTGAGCTTGAAGAGAAAGAAAGTGAATATGCTTACCTAGCCTAGACAGACTTGG	1735
Db	124108	ACTGTGAGCTTGAAGAGAAAGAAAGTGAATATGCTTACCTAGCCTAGACAGACTTGG	1240490
Oy	1736	GATGGTCCAGGGCTCTATGTGACCTTCAAGACAAAGAAAGACTTCGACAGTCTAGGTTC	1795
Db	124048	GATGGTCCAGGGCTCTATGTGACCTTCAAGACAAAGAAAGACTTCGACAGTCTAGGTTC	1239890
Oy	1796	CTCAAAATGCCCCCAATTGAAGACACAGGCCCACTCTTTTCTTTTCTTTTGTAGAAGGA	1855
Db	123988	CTCAAAATGCCCCCAATTGAAGACACAGGCCCACTCTTTTCTTTTCTTTTGTAGAAGGA	1239390
Oy	1856	GTCCTGCTGTGGTCCCACTGAGTGCATGGCAGATCTCAGCTCACTGCACCTCC	1915

Db 123929	GTCTTGCCCTGGTTGGCCATCTGTAAGTGCATATGGACAGATCTCAGCTCACTGCAACTCC	123870
Qy 1916	ATCTCTCGAGATTCAACAAATTCCGCGCTCAGCTCCAGAAATGCTGGAGATTACAGGCG	19175
Db 123869	ATCTCTCGAGATTCAACAAATTCCGCGCTCAGCTCCAGAAATGCTGGAGATTACAGGCG	123810
Qy 1976	TACACACACATGCTCGCTCAATTTTTTTTGTATTTTTTAAGTAAATGGGGTTCCACACAT	2035
Db 123809	TACACACACATGCTCGCTCAATTTTTTTTGTATTTTTTAAGTAAATGGGGTTCCACACAT	123750
Qy 2036	TGGCCAGGCTGGTGTGCACACTCCTGACCTCAGGTGATCCACCCACCTTGACCTCCCAAG	2095
Db 123749	TGGCCAGGCTGGTGTGCACACTCCTGACCTCAGGTGATCCACCCACCTTGACCTCCCAAG	123690
Qy 2096	TGCTGGGATTACAGGTGTGAGCAGGAGCAGGACCACTTCTCAGATCTCAATTTCATT	2155
Db 123689	TGCTGGGATTACAGGTGTGAGCAGGAGCAGGACCACTTCTCAGATCTCAATTTCATT	123630
Qy 2156	TTGTGGCTTACCAATTCCTTACGACACTGCGCTTGCCATTTGTGTGGCGAATATAAAAATA	2215
Db 123629	TTGTGGCTTACCAATTCCTTACGACACTGCGCTTGCCATTTGTGTGGCGAATATAAAAATA	123570
Qy 2216	CACCTTAAAGTCTAGACACACTCAGTAGAGCCAGCACCCTCAGTGCTGGGACGGGCGAT	2275
Db 123669	CACCTTAAAGTCTAGACACACTCAGTAGAGCCAGCACCCTCAGTGCTGGGACGGGCGAT	123510
Qy 2276	CAGAAAGTGTAAAGCCCTCTCTCCACATGTGCAGAAAGGAGACACAGCAGCTCACAACAT	2335
Db 123509	CAGAAAGTGTAAAGCCCTCTCTCCACATGTGCAGAAAGGAGACACAGCAGCTCACAACAT	123450
Qy 2336	CCAGGCTTGAATTTCCCTGCTCTCATATAACAGAAAGAGTCTGTGGATCCGCTTAAG	2395
Db 123449	CCAGGCTTGAATTTCCCTGCTCTCATATAACAGAAAGAGTCTGTGGATCCGCTTAAG	123390
Qy 2396	GGATCAGGAGAGGAGAAAGAAAGAGATGGGGTGGGAGGACCCCTCCAGTCTCTTACT	2455
Db 123389	GGATCAGGAGAGGAGAAAGAAAGAGATGGGGTGGGAGGACCCCTCCAGTCTCTTACT	123330
Qy 2456	GGTTTCCCAACTACAGGTGGGGGTGGGAAAGGCTTATCAGATATCAACAGGTTTCA	2515
Db 123329	GGTTTCCCAACTACAGGTGGGGGTGGGAAAGGCTTATCAGATATCAACAGGTTTCA	123270
Qy 2516	ATTAAAGATTGATTTATTCAGATGTGAAAAAATTTTTTTTTTTTTTTT	2564
Db 123269	ATTAAAGATTGATTTATTCAGATGTGAAAAAATTTTTTTTTTTTTTTT	123221

RESULT 4	AX443133	1183 bp	DNA	linear	PAT 02-JUL-2002
LOCUS	AX443133				
DEFINITION	Sequence 74 from Patent WO0216599.				
ACCESSION	AX443133				
VERSION	AX443133.1	GI:21690555			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1				
	Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shinkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.				
TITLE	Proteins and nucleic acids encoding same				
JOURNAL	Patent: WO 0216599-A 74 28-FEB-2002;				
	Cutagen Corporation (US); COR THERAPEUTICS, INC. (US)				
FEATURES	Location/Qualifiers				
source	1..1183				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	251 a 359 c 333 g 240 t				
ORIGIN					

Query Match 46.0%; Score 1181.4; DB 6; Length 1183;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 18 AGCTAGAGCTCCAGAGACCCACGCTGTGTCTGTGACAGAGCTCAAAGGCTCGG 77
 Db 1 AGCTAGAGCTCCAGAGACCCACGCTGTGTCTGTGACAGAGCTCAAAGGCTCGG 60

Qy 78 CTTTCCCTCCCTGGCTGGCTGTGTGAGAGGTTCCCAAGTCCCAAGTCCCTTAAGAG 137
 Db 61 CTTTCCCTCCCTGGCTGGCTGTGTGAGAGGTTCCCAAGTCCCTTAAGAG 120

Qy 138 CATGGGACAGTGAATCCATCTGTGTGAACACTGTGACTGACAGAGATGTGAGCT 197
 Db 121 CATGGGACAGTGAATCCATCTGTGTGAACACTGTGACTGACAGAGATGTGAGCT 180

Qy 198 ACCCAACCAACACTAGCTCTCCCTGAAATCTCCAGGCTGAGAGATTTGGGTTG 257
 Db 181 ACCCAACCAACACTAGCTCTCCCTGAAATCTCCAGGCTGAGAGATTTGGGTTG 240

Qy 258 TCTTAGACCAAGAGACATGGCACTTCCAGAAAGGCCCCCAAGCCCTTAAGCTGTCA 317
 Db 241 TCTTAGACCAAGAGACATGGCACTTCCAGAAAGGCCCCCAAGCCCTTAAGCTGTCA 300

Qy 318 GCCAGAGCATGGCTCTCAGAGAGCTGTCTCCCAAGCTTTGATGACAAACCAATTC 377
 Db 301 GCCAGAGCATGGCTCTCAGAGAGCTGTCTCCCAAGCTTTGATGACAAACCAATTC 360

Qy 378 CTCGATGATGTGCTTCTGAGTGTCTGTGAGAGCAATGGAGCTGCCACAGAGAG 437
 Db 361 CTCGATGATGTGCTTCTGAGTGTCTGTGAGAGCAATGGAGCTGCCACAGAGAG 420

Qy 438 AAAATCTCTGCAAGCCCAAGCTGATGCTCTGTCTCAAGGCAAGGAGCTGTGACAT 497
 Db 421 AAAATCTCTGCAAGCCCAAGCTGATGCTCTGTCTCAAGGCAAGGAGCTGTGACAT 480

Qy 498 GGAAGCAAGAGAGCAAGGCAAGCCGTGGCCCTTGGGAGTTTCCCGGAGGTGGCCC 557
 Db 481 GGAAGCAAGAGAGCAAGGCAAGCCGTGGCCCTTGGGAGTTTCCCGGAGGTGGCCC 540

Qy 558 GGCAGAGCTGTCTGTGAGACTGGGGAGACATTTGACATCTGTCTGTGAGATGAGAC 617
 Db 541 GGCAGAGCTGTCTGTGAGACTGGGGAGACATTTGACATCTGTCTGTGAGATGAGAC 600

Qy 618 GTGAGAGGTGTCTGTGAAGTCTCAGGAGAGAGATTAACATCCACGTCACAGTGC 677
 Db 601 GTGAGAGGTGTCTGTGAAGTCTCAGGAGAGAGATTAACATCCACGTCACAGTGC 660

Qy 678 CAAAGTCTCCCATGGGTGGCTGTATGAGAGGCTGAGAGAGAGAGAGAGAGAGTGT 737
 Db 661 CAAAGTCTCCCATGGGTGGCTGTATGAGAGGCTGAGAGAGAGAGAGAGAGTGT 720

Qy 738 GTTGTTCCTTGGAACTCTGAGAGGGCTTCTGTATCCGGAAGCCAGACAGAGAGAG 797
 Db 721 GTTGTTCCTTGGAACTCTGAGAGGGCTTCTGTATCCGGAAGCCAGACAGAGAGAG 780

Qy 798 CTCTTACTCTGTCAAGTCCGCTCAGAGCCGCTGAGATCTGTGAGAGCCGAGAGACT 857
 Db 781 CTCTTACTCTGTCAAGTCCGCTCAGAGCCGCTGAGATCTGTGAGAGCCGAGAGACT 840

Qy 858 CAGATTCACATGCTTGAACATGAGGCTGTGACATTCACCGGCTCACTTCCCTC 917
 Db 841 CAGATTCACATGCTTGAACATGAGGCTGTGACATTCACCGGCTCACTTCCCTC 900

Qy 918 ACTCAGAGCCCTGTGAGACCATTAATCTGAGAGGAGAGAGATCTGTGAGAGAGAG 977
 Db 901 ACTCAGAGCCCTGTGAGACCATTAATCTGAGAGGAGAGAGATCTGTGAGAGAGAG 960

Qy 978 GGAAGCTGTGTCTGTGAGAGGCTGGCCGCTCTCCGGAAGAGATATACCTTACTGT 1037
 Db 961 GGAAGCTGTGTCTGTGAGAGGCTGGCCGCTCTCCGGAAGAGATATACCTTACTGT 1020

Qy 1038 GACTGTGAGAGACCACTCACTGAGAAAGAGCTGAGACAGCTCCCTCTGTTTCTGA 1097

Db 1021 GACTGTGAGAGACCACTCACTGAGAAAGAGCTGAGACAGCTCCCTGTTTCTGA 1080

Qy 1098 AGCTGCAACAGGAGAGAGCTCTTCTGATGAGGCTCTCGGAGAGCTCCAGCTTCA 1157
 Db 1081 AGCTGCAACAGGAGAGAGCTCTTCTGATGAGGCTCTCGGAGAGCTCCAGCTTCA 1140

Qy 1158 CATCAGCTGATGATGAGAGCTGTCTCTTTGATGATGAGCTTGA 1200
 Db 1141 CATCAGCTGATGATGAGAGCTGTCTCTTTGATGATGAGCTTGA 1183

RESULT 5
 AX443135/c
 LOCUS AX443135 1183 bp DNA linear PAT 02-JUL-2002
 DEFINITION Sequence 76 from Patent WO0216599.
 ACCESSION AX443135
 VERSION AX443135.1 GI:21690556
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Burgess, C.E., Conley, P.B., Grosse, W.M., Hart, M., Kekuda, R.,
 Shimkets, R.A., Splet, K.A., Szekeres, E.S., Tomlinson, J.E.,
 Topper, J.N. and Yang, R.B.
 Proteins and nucleic acids encoding same
 Patent: WO 0216599-A 76 28 FEB-2002;
 Curagen Corporation (US); COR THERAPEUTICS, INC. (US)

TITLE
 JOURNAL
 LOCATION/Qualifiers

FEATURES
 source
 1..1183
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 240 a 333 c 359 g 251 t

ORIGIN

Query Match 46.0%; Score 1181.4; DB 6; Length 1183;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1182; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 18 AGCTAGAGCTCCAGAGACCCACGCTGTGTCTGTGACAGAGCTCAAAGGCTCGG 77
 Db 1183 AGCTAGAGCTCCAGAGACCCACGCTGTGTCTGTGACAGAGCTCAAAGGCTCGG 1124

Qy 78 CTTTCCCTCCCTGGCTGGCTGTGTGAGAGGTTCCCAAGTCCCTTAAGAGAG 137
 Db 1123 CTTTCCCTCCCTGGCTGGCTGTGTGAGAGGTTCCCAAGTCCCTTAAGAGAG 1064

Qy 138 CATGGGACAGTGAATCCATCTGTGTGAACACTGTGACTGACAGAGATGTGAGCT 197
 Db 1063 CATGGGACAGTGAATCCATCTGTGTGAACACTGTGACTGACAGAGATGTGAGCT 1004

Qy 198 ACCCAACCAACACTAGCTCTCCCTGAAATCTCCAGGCTGAGAGATTTGGGTTG 257
 Db 1003 ACCCAACCAACACTAGCTCTCTCTGAAATCTCCAGGCTGAGAGATTTGGGTTG 944

Qy 258 TCTTAGACCAAGAGACATGGCACTTCCAGAAAGGCCCCCAAGCCCTTAAGCTGTCA 317
 Db 943 TCTTAGACCAAGAGACATGGCACTTCCAGAAAGGCCCCCAAGCCCTTAAGCTGTCA 884

Qy 318 GCCAGAGCATGGCTCTCAGAGAGCTGTCTCCCAAGCTTTGATGACAAACCAATTC 377
 Db 883 GCCAGAGCATGGCTCTCAGAGAGCTGTCTCCCAAGCTTTGATGACAAACCAATTC 824

Qy 378 CTCGATGATGTGCTTCTGAGTGTCTGTGAGAGAACTATGGAGAGTGTGCCAGAGAG 437
 Db 823 CTCGATGATGTGCTTCTGAGTGTCTGTGAGAGAACTATGGAGAGTGTGCCAGAGAG 764

Qy 438 AAAATCTCTGCAAGCCCAAGCTGATGCTCTGTCTCAAGGCAAGGAGCTGTGACAT 497
 Db 763 AAAATCTCTGCAAGCCCAAGCTGATGCTCTGTCTCAAGGCAAGGAGCTGTGACAT 704

Grand-Pierre, N., Grant, G., Hagos, B., Heaton, A., Johnson, L., Kaitas, A., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Kartas, T., Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McCarthy, M., McEwen, P., McGuck, A., McKernan, K., McSheeters, R., Melchior, J., Menses, L., Minova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, C., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfayev, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, D., Vestiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced g1:Tig81243.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L7115
Center clone name: 712.N.14

Summary Statistics
Sequencing vector: M13, M7815, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 125577 bases at least Q40
Consensus quality: 135703 bases at least Q30
Consensus quality: 139593 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 142233; sum-of-contigs
Quality coverage: 2.6 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 37 contigs. The true order of the pieces * is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will be preserved.

1	1214:	contig of 1214 bp in length
*	1315	1314: gap of 100 bp
*	1315	2673: contig of 1359 bp in length
*	2674	2773: gap of 100 bp
*	2774	4520: contig of 1747 bp in length
*	4521	4620: gap of 100 bp
*	4621	5961: contig of 1341 bp in length
*	5962	6061: gap of 100 bp
*	6062	7719: contig of 1658 bp in length
*	7720	7819: gap of 100 bp
*	7820	9793: contig of 1980 bp in length
*	9800	9899: gap of 100 bp
*	9900	11434: contig of 1535 bp in length
*	11435	11534: gap of 100 bp
*	11535	14382: contig of 2848 bp in length
*	14383	14482: gap of 100 bp
*	14483	17148: contig of 2666 bp in length
*	17149	17248: gap of 100 bp
*	17249	18032: contig of 1891 bp in length

NOTE: This is a 'working draft' sequence. It currently consists of 37 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1214:	contig of 1214 bp	in length
*	1215	1314:	gap of 100 bp
*	1315	2673:	contig of 1356 bp in length
*	2674	2773:	gap of 100 bp
*	2774	4520:	contig of 1747 bp in length
*	4521	4620:	gap of 100 bp
*	4621	5961:	contig of 1341 bp in length
*	5962	6061:	gap of 100 bp
*	6062	7719:	contig of 1658 bp in length
*	7720	7819:	gap of 100 bp
*	7820	9799:	contig of 1980 bp in length
*	9800	9899:	gap of 100 bp
*	9900	11433:	contig of 1535 bp in length
*	11435	11534:	gap of 100 bp
*	11535	14382:	contig of 2846 bp in length
*	14383	14482:	gap of 100 bp
*	14483	17148:	contig of 2666 bp in length
*	17149	17248:	gap of 100 bp
*	17249	19233:	contig of 1991 bp in length
*	19240	19339:	gap of 100 bp
*	19340	21107:	contig of 1763 bp in length
*	21103	21202:	gap of 100 bp
*	21203	23377:	contig of 2165 bp in length
*	23372	234771:	gap of 100 bp

MEDLINE 21553259
 PUBMED 11696592
 REFERENCE 2 (bases 1 to 786)
 AUTHORS Holland, S. J., Mendenhall, M. K., Zhou, X., Spencer, C., Pardo, J.,
 Fu, A. C., Sheng, N., Shen, M., Liao, C., Luo, Y., Payan, D. G.,
 Manebo, H. S. Y., and Wu, J.

TITLE Direct Submission
 JOURNAL Submitted (05-DEC-2000) Rigel Pharmaceutical Inc., 240 East Grand
 Avenue, South San Francisco, CA 94080, USA
 FEATURES Location/Qualifiers
 source 1..786
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="20"
 1..786
 /note="SLP2; SH2 and SH3 domain-containing adaptor"
 /codon_start=1
 /product="Src-like adaptor protein-2"
 /protein_id="AA129204.1"
 /db_xref="GI:16797892"
 /translation="MGSLSRKRSLPSLSVSGVGGPYTMEARSKATVALGSPF
 AGPAELSLRLGEPLTIVSEBDQMTVLSVSGREYNIPIVHAKVSHGMLYEGLSRE
 KAEELILPQNGAFILRESQTRGYSLSVLSRPAWDRIHRYRIHCLDNGMLYI
 SPRLTFPSLOALVDHYSLEADICCLKEPCVLOAGPLGKDIPLPTVQRTPLNWK
 ELDSLLFEEATGSESLSEGLRESLSFYISLNDENVSLDA"

CDS

BASE COUNT 162 a 234 c 231 g 159 t
 ORIGIN
 Query Match 30.6%; Score 786; DB 9; Length 786;
 Best Local Similarity 100.0%; Pred. No. 1.9e-217;
 Matches 786; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 415 ATGGAGAGTCTGCCACAGAGAAATCTCTGCCAGCCCAAGCTGATGTTCTCTGTC 474
 Db 1 ATGGAGAGTCTGCCACAGAGAAATCTCTGCCAGCCCAAGCTGATGTTCTCTGTC 60
 QY 475 CAAGGCCAGGAGCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534
 Db 61 CAAGGCCAGGAGCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 535 GGCAGTTTCCGGCAGGTGGCCCGGCGCAGCTGCTGCTGAGAGCTGGGGAGCATTTGACC 594
 Db 121 GGCAGTTTCCGGCAGGTGGCCCGGCGCAGCTGCTGCTGAGAGCTGGGGAGCATTTGACC 180
 QY 595 ATCGTCTCTGAGATGAGAGCTGTGTGACAGGTGCTGTGAACTCTCAGGACAGAGATAT 654
 Db 181 ATCGTCTCTGAGATGAGAGCTGTGTGACAGGTGCTGTGAACTCTCAGGACAGAGATAT 240
 QY 655 AACATCCCAAGCTCCAGTGGCCAAAGTCTCCCATGGGTGGTGTATGAGGGCTGAGAC 714
 Db 241 AACATCCCAAGCTCCAGTGGCCAAAGTCTCCCATGGGTGGTGTATGAGGGCTGAGAC 300
 QY 715 AGGAG 774
 Db 301 AGGAG 360
 QY 775 CGGAGAGCCAGACCAAGAGAGAGCTTACTCTGTGAGTCCGCTCAGCCGCTGCA 834
 Db 361 CGGAGAGCCAGACCAAGAGAGAGCTTACTCTGTGAGTCCGCTCAGCCGCTGCA 420
 QY 835 TCCTGGAGCCGGATCAGACATTCAGAGATTCATGCTGTTGACAAATGCTGCTGATC 894
 Db 421 TCCTGGAGCCGGATCAGACATTCAGAGATTCATGCTGTTGACAAATGCTGCTGATC 480
 QY 895 TCACCGGCGCTCACTTCCCTCACTCCAGGCTTGGTGGACATTAATCTGAGCTGGG 954
 Db 481 TCACCGGCGCTCACTTCCCTCACTCCAGGCTTGGTGGACATTAATCTGAGCTGGG 540
 QY 955 GATGACATCTGCTGCTCAAGAGAGCCGTGTCTCTGAGAGAGAGAGAGAGAGAGAGAG 1014
 Db 541 GATGACATCTGCTGCTCAAGAGAGCCGTGTCTCTGAGAGAGAGAGAGAGAGAGAGAG 600
 QY 1015 GGCAGAGATATATCCCTCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074

Db 601 GGCAGAGATATATCCCTCACTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 QY 1075 GACAGCTCCCTCTGTTTCTGAAGCTCCACAGGAGAGAGAGAGAGAGAGAGAGAGAG 1134
 Db 661 GACAGCTCCCTCTGTTTCTGAAGCTCCACAGGAGAGAGAGAGAGAGAGAGAGAGAG 720
 QY 1135 CTCGGGAGTCCCTCAGCTTCTTACATCAGCTGATGACGAGGCTGTCTTTGATGAT 1194
 Db 721 CTCGGGAGTCCCTCAGCTTCTTACATCAGCTGATGACGAGGCTGTCTTTGATGAT 780
 QY 1195 GCCTAG 1200
 Db 781 GCCTAG 786

RESULT 9

AF290986 737 bp mRNA linear PRI 22-APR-2002
 LOCUS Homo sapiens Src-like adaptor protein-2 splice isoform mRNA,
 DEFINITION complete cds.
 ACCESSION AF290986
 VERSION AF290986.1 GI:17351922
 KEYWORDS

SOURCE

ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 737)
 Loreto, M.P. and McGlade, C.J.
 TITLE Direct Submission
 JOURNAL Submitted (28-JUL-2000) Brain Tumour Research Centre, Hospital for
 Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada

FEATURES

source 1..737
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="20"
 /ciseq_type="thymus"
 1..633
 /note="SLP-2-v"
 /codon_start=1
 /product="Src-like adaptor protein-2 splice isoform"
 /protein_id="AA138198.1"
 /db_xref="GI:17351923"
 /translation="MGSLSRKRSLPSLSVSGVGGPYTMEARSKATVALGSPF
 AGPAELSLRLGEPLTIVSEBDQMTVLSVSGREYNIPIVHAKVSHGMLYEGLSRE
 KAEELILPQNGAFILRESQTRGYSLSVLSRPAWDRIHRYRIHCLDNGMLYI
 SPRLTFPSLOALVDHYSLEADICCLKEPCVLOAGPLGKDIPLPTVQRTPLNWK
 ELDSLLFEEATGSESLSEGLRESLSFYISLNDENVSLDA"

CDS

BASE COUNT 152 a 219 c 218 g 148 t
 ORIGIN
 Query Match 25.8%; Score 663.4; DB 9; Length 737;
 Best Local Similarity 93.4%; Pred. No. 9.8e-182;
 Matches 735; Conservative 0; Mismatches 1; Indels 51; Gaps 2;

QY 415 ATGGAGAGTCTGCCACAGAGAAATCTCTGCCAGCCCAAGCTGATGTTCTCTGTC 474
 Db 1 ATGGAGAGTCTGCCACAGAGAAATCTCTGCCAGCCCAAGCTGATGTTCTCTGTC 60
 QY 475 CAAGGCCAGGAGCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534
 Db 61 CAAGGCCAGGAGCTGTGACCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 535 GGCAGTTTCCGGCAGGTGGCCCGGCGCAGCTGCTGCTGAGAGCTGGGGAGCATTTGACC 594
 Db 121 GGCAGTTTCCGGCAGGTGGCCCGGCGCAGCTGCTGCTGAGAGCTGGGGAGCATTTGACC 180
 QY 595 ATCGTCTCTGAGATGAGAGCTGTGTGACAGGTGCTGTGAACTCTCAGGACAGAGATAT 654
 Db 181 ATCGTCTCTGAGATGAGAGCTGTGTGACAGGTGCTGTGAACTCTCAGGACAGAGATAT 240
 QY 655 AACATCCCAAGCTCCAGTGGCCAAAGTCTCCCATGGGTGGTGTATGAGGGCTGAGAC 714

Quality coverage: 3.3 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 37 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 1214: contig of 1214 bp in length
1215 1314: gap of 100 bp
1315 2673: contig of 1359 bp in length
2674 2773: gap of 100 bp
2774 4520: contig of 1747 bp in length
4521 4620: gap of 100 bp
4621 5961: contig of 1341 bp in length
5962 6061: gap of 100 bp
6062 7719: contig of 1658 bp in length
7720 9759: contig of 1980 bp in length
9800 9899: gap of 100 bp
9900 11434: contig of 1535 bp in length
11435 11535: gap of 100 bp
11535 14382: contig of 2848 bp in length
14383 14482: gap of 100 bp
14483 17148: contig of 2666 bp in length
17149 17248: gap of 100 bp
17249 19239: contig of 1991 bp in length
19240 19339: gap of 100 bp
19340 21102: contig of 1763 bp in length
21103 21202: gap of 100 bp
21203 23371: contig of 2169 bp in length
23372 23471: gap of 100 bp
23472 25782: contig of 2311 bp in length
25783 25882: gap of 100 bp
25883 28824: contig of 2942 bp in length
28825 28924: gap of 100 bp
28925 31619: contig of 2695 bp in length
31620 31719: gap of 100 bp
31720 34679: contig of 2960 bp in length
34680 34779: gap of 100 bp
34780 37527: contig of 2748 bp in length
37528 37627: gap of 100 bp
37628 40146: contig of 2519 bp in length
40147 40246: gap of 100 bp
40247 43743: contig of 3497 bp in length
43744 43843: gap of 100 bp
43844 47133: contig of 3280 bp in length
47134 47233: gap of 100 bp
47234 51023: contig of 3790 bp in length
51024 51123: gap of 100 bp
51124 54935: contig of 3812 bp in length
54936 55035: gap of 100 bp
55036 59553: contig of 4518 bp in length
59554 59653: gap of 100 bp
59654 64428: contig of 4775 bp in length
64429 64528: gap of 100 bp
64529 69211: contig of 4683 bp in length
69212 69311: gap of 100 bp
69312 72901: contig of 3590 bp in length
72902 73001: gap of 100 bp
73002 76724: contig of 3723 bp in length
76725 76824: gap of 100 bp
76825 81179: contig of 4355 bp in length
81180 81279: gap of 100 bp
81280 87009: contig of 5730 bp in length
87010 87109: gap of 100 bp
87110 90855: contig of 3746 bp in length
90856 90955: gap of 100 bp
90956 96520: contig of 5565 bp in length
96521 96620: gap of 100 bp
96621 102321: contig of 5701 bp in length
```

FEATURES

source

```
* 102322 102421: gap of 100 bp
* 102422 108293: contig of 5872 bp in length
* 108294 108393: gap of 100 bp
* 108394 116689: contig of 8286 bp in length
* 116690 116789: gap of 100 bp
* 116790 125264: contig of 8475 bp in length
* 125265 125364: gap of 100 bp
* 125365 136354: contig of 10990 bp in length
* 136355 136454: gap of 100 bp
* 136455 145833: contig of 9379 bp in length.
```

Location/Qualifiers
1. .145833
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="20"
/map="20"

```
/clone="RP11-712N14"  
1. .1214  
/clone_lib="RP11-11 Human Male BAC"  
misc_feature  
1315. .2673  
/note="assembly_fragment"  
misc_feature  
2774. .4520  
/note="assembly_fragment"  
misc_feature  
4621. .5961  
/note="assembly_fragment"  
misc_feature  
6062. .7719  
/note="assembly_fragment"  
7820. .9799  
/note="assembly_fragment"  
9900. .11434  
/note="assembly_fragment"  
11535. .14382  
/note="assembly_fragment"  
14483. .17148  
/note="assembly_fragment"  
17249. .19239  
/note="assembly_fragment"  
19340. .21102  
/note="assembly_fragment"  
21203. .23371  
/note="assembly_fragment"  
23472. .25782  
/note="assembly_fragment"  
25783. .25882  
/note="assembly_fragment"  
28825. .28924  
/note="assembly_fragment"  
31620. .31719  
/note="assembly_fragment"  
34680. .34779  
/note="assembly_fragment"  
37528. .37627  
/note="assembly_fragment"  
40147. .40246  
/note="assembly_fragment"  
43744. .43843  
/note="assembly_fragment"  
47134. .47233  
/note="assembly_fragment"  
51024. .51123  
/note="assembly_fragment"  
54936. .55035  
/note="assembly_fragment"  
59554. .59653  
/note="assembly_fragment"  
64429. .64528  
/note="assembly_fragment"  
69212. .69311  
/note="assembly_fragment"  
72902. .73001  
/note="assembly_fragment"  
76725. .76824  
/note="assembly_fragment"  
81180. .81279  
/note="assembly_fragment"  
87010. .87109  
/note="assembly_fragment"  
90856. .90955  
/note="assembly_fragment"  
96521. .96620  
/note="assembly_fragment"
```

misc_feature

vector_end:T7
vector_side:right
/note="assembly_fragment"

Query March 14.1k: Score 362; DB 2; Length 145833;
Best Local Similarity 96.1k: Pred. No. 1.2e-33;
Matches 371; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 9 TCCGTCGAGCTAAGCTCCAGAGCCACGCTGTCTCTGTGACAGCTCAAG 68

	Db	99736	TGACGTCCGAAGCTTAGAGCTCAAGAAGCCACGACTGTGTCTGTGACAAGTCAAG	99795
REFERENCE	Oy	69	GCCCCGGGCCCTTCCTCCCTCGCTGGCTGTGCTTTGGAGAGGTTCCCACTCAGATC	128
REFERENCE	Db	99796	GCCCCGAGGCTTCTCCCTCCGTGGCTGGCTGTGGAGGGTTCCCACTCAGAAATC	99855
REFERENCE	Oy	129	CCTAAGAGAGCATGGGGCAGCTGATTCATCCCTCGTGTAACAACTGCATGACAGACAA	188
REFERENCE	Db	99856	CCTAAGAGAGCATGGGGCAGCTGATTCATCCCTCGTGTAACAACTGCATGACAGACAA	99915
REFERENCE	Oy	189	TGCTAGAGTACCACAAACCAAACCTAGCTTTCCTGAAATCTTCCAGGCTGAGAGAG	248
REFERENCE	Db	99916	TGCTAGAGTACCACAAACCAAACCTAGCTTTCCTGAAATCTTCCAGGCTGAGAGAG	99975
REFERENCE	Oy	249	TTCTGGGTGTCCTTAGAGCAAGAGACATGGCAGACTTCCAGAAAGGCCCCCAAGGCCCTA	308
REFERENCE	Db	99976	TTCTGGGTGTCCTTAGAGCAAGAGACATGGCAGACTTCCAGAAAGGCCCCCAAGGCCCTA	100035
REFERENCE	Oy	309	ACCTGTCCAGCAGAGACATGGGTGTACAGAGAGCTGCTTCCCAAGCTTTATGACAAA	368
REFERENCE	Db	100036	ACCTGTCCAGCAGAGACATGGGTGTACAGAGAGCTGCTTCCCAAGCTTTATGACAAA	100095
REFERENCE	Oy	369	CCAAATTTCCTCCATGATGTGCTTCT	394
REFERENCE	Db	100096	CCAGTAAGTATGCTTGTCTTTGTCT	100121
RESULT 14				
LOCUS	AC021778	147595 bp	DNA	linear HTG 12-MAR-2000
DEFINITION	Homo sapiens clone RP11-14D1, WORKING DRAFT SEQUENCE, 24 unordered pieces.			
ACCESSION	AC021778			
VERSION	AC021778.3	GI:7230095		
KEYWORDS	HTG; HTGS PHASE1; HTGS_DRAFT.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Bairren, B., Linton, L., Nusbäum, C. and Lander, E.			
JOURNAL	Homo sapiens, clone RP11-14D1			
REFERENCE	Unpublished			
AUTHORS	2 (bases 1 to 147595)			
	Bairren, B., Linton, L., Nusbäum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barina, N., Beckert, R., Bedalov, F., Boguski, M.S., Bouknight, B., Brown, A., Burkett, G., Castle, A., Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeRubeis, K., Dewar, K., Domingo, M., Doyle, M., Penetier, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, R., Heaford, A., Horton, L., Howland, J.C., Johnson, R., Jones, C., Kam, I., Karasik, A., Klein, J., Landers, T., Lechick, J., Levine, R., Lien, C., Liu, G., Locke, K., McDonald, P., Margulis, N., McEwan, P., McGurk, A., McKernan, K., McPherson, R., Meldrum, J., Meneses, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K., Pierre, N., Pisanti, C., Pollard, V., Raymond, C., Riley, R., Roach, D., Roy, A.L., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Scjnovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Trielli, A., Vasiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W., Zimer, A. and Zody, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (20-JUN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
COMMENT	On Mar 12, 2000 this sequence version replaced gi:6980933. All repeats were identified using RepeatMasker:			
	Smith, A.F.A. & Green, P. (1996-1997)			
	http://ftp.genome.washington.edu/SW/RepeatMasker.html			
	----- Genome Center			
	Center: Whitehead Institute/ MIT Center for Genome Research			
	Center code: WIBB			
	Web site: http://www-seq.wi.mit.edu			
	Contact: sequence_submissions@genome.wi.mit.edu			

```

Project Information
Center project name: L3390
Center clone name: 14 D 1
Summary Statistics
Sequencing vector: M13, M7815, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap, version 0.960731
Consensus quality: 138993 bases at least Q40
Consensus quality: 143079 bases at least Q30
Consensus quality: 144508 bases at least Q20
Insert size: 153000; agarose-fp
Insert size: 145295; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
consists of 24 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1
1103: contig of 1103 bp in length
1104 1203: gap of 100 bp
1204 2484: contig of 1281 bp in length
2485 2584: gap of 100 bp
2585 3491: contig of 907 bp in length
3492 3591: gap of 100 bp
3592 5191: contig of 1600 bp in length
5192 5291: gap of 100 bp
5292 6480: contig of 1189 bp in length
6481 6580: gap of 100 bp
6581 10605: contig of 4025 bp in length
10606 10705: gap of 100 bp
10706 15069: contig of 4364 bp in length
15070 15159: gap of 100 bp
15170 19127: contig of 3958 bp in length
19128 19227: gap of 100 bp
19228 23443: contig of 4216 bp in length
23444 23543: gap of 100 bp
23544 27726: contig of 4183 bp in length
27727 27826: gap of 100 bp
27827 32101: contig of 4275 bp in length
32102 32201: gap of 100 bp
32202 38019: contig of 5818 bp in length
38020 38119: gap of 100 bp
38120 44567: contig of 6448 bp in length
44568 44667: gap of 100 bp
44668 49571: contig of 4904 bp in length
49572 49671: gap of 100 bp
49672 55655: contig of 5984 bp in length
55656 55755: gap of 100 bp
55756 62547: contig of 6792 bp in length
62548 62647: gap of 100 bp
62648 69871: contig of 7224 bp in length
69872 69971: gap of 100 bp
69972 78995: contig of 9024 bp in length
78996 79095: gap of 100 bp
79096 87813: contig of 8718 bp in length
87814 87913: gap of 100 bp
87914 96855: contig of 8942 bp in length
96856 96955: gap of 100 bp
96956 107916: contig of 10961 bp in length
107917 108015: gap of 100 bp
108017 118688: contig of 10672 bp in length
118689 118788: gap of 100 bp
118789 132452: contig of 13664 bp in length
132453 132552: gap of 100 bp
132553 147595: contig of 15043 bp in length.

Location/Qualifiers
1. 147595
/organism="Homo sapiens"
source

```

/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="RP11-14D1"
/clone_1fb="RP11-11 Human Male BAC"
1..1103
misc_feature
  /note="assembly_fragment"
  1204..2484
misc_feature
  /note="assembly_fragment"
  2585..3491
misc_feature
  /note="assembly_fragment"
  clone_end:SP6
  vector_side:right"
3592..5191
misc_feature
  /note="assembly_fragment"
  5292..6480
misc_feature
  /note="assembly_fragment"
  6581..10605
misc_feature
  /note="assembly_fragment"
  10706..15069
misc_feature
  /note="assembly_fragment"
  15170..19127
misc_feature
  /note="assembly_fragment"
  19228..23443
misc_feature
  /note="assembly_fragment"
  23544..27726
misc_feature
  /note="assembly_fragment"
  27827..32101
misc_feature
  /note="assembly_fragment"
  32202..38019
misc_feature
  /note="assembly_fragment"
  38120..44567
misc_feature
  /note="assembly_fragment"
  44668..49571
misc_feature
  /note="assembly_fragment"
  49672..55655
misc_feature
  /note="assembly_fragment"
  55756..62547
misc_feature
  /note="assembly_fragment"
  62648..69871
misc_feature
  /note="assembly_fragment"
  clone_end:T7
  vector_side:right"
69972..78995
misc_feature
  /note="assembly_fragment"
  79096..87813
misc_feature
  /note="assembly_fragment"
  87914..96855
misc_feature
  /note="assembly_fragment"
  96956..107916
misc_feature
  /note="assembly_fragment"
  108017..118688
misc_feature
  /note="assembly_fragment"
  118789..132452
misc_feature
  /note="assembly_fragment"
  132553..147595
misc_feature
  /note="assembly_fragment"

BASE COUNT  40083 a 32356 c 32238 g 40610 t 2308 others
ORIGIN
Query Match      9.8%; Score 251.2; DB 2; Length 147595;
Best Local Similarity 80.2%; Pred. No. 2.3e-61;
Matches 295; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

```

```

QY 1958 TACCTGGATTACAGGCGTACACCACTGCGGTGCTAATTTTGTATTTTGTAGTA 2017
DB 8418 TACCTGGATTATATGAGGCGTACACCACTGCGGTGCTAATTTTGTATTTTGTAGTA 8477
QY 2018 CATGGGTTTACACCATTTGGCCAGGCTGTGTGAATCTTACCTCAGGTATCCACC 2077
DB 8478 GACAGGTTTACACCATTTGGCCAGGCTGTGTGAATCTTACCTCAGGTATCCACC 8537
QY 2078 CACCTTGGCTTCCCAAGGCTGTGTGAATTTACAGGTGAGCAAGGCAAGGCTGAGCC 2137
DB 8538 CACCTTGGCTTCCCAAGGCTGTGTGAATTTACAGGTGAGCAAGGCAAGGCTGAGCC 8597
QY 2138 TCAGATCT 2145
DB 8598 TAACACT 8605

```

```

RESULT 15
AC099066/c 149548 bp DNA linear pri 27-FEB-2002
LOCUS
DEFINITION Homo sapiens chromosome 1 clone RP11-145A3, complete sequence.
ACCESSION AC099066 AL59830
VERSION AC099066.2 GI:18958741
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 149548)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submission
JOURNAL
  2 (bases 1 to 149548)
  Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
  Direct Submission
  Submitted (08-NOV-2001) Genome Center, University of Washington,
  Box 352145, Seattle, WA 98195, USA
  3 (bases 1 to 149548)
  Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
  Saenphimachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
  Direct Submission
  Submitted (27-FEB-2002) Genome Center, University of Washington,
  Box 352145, Seattle, WA 98195, USA
  On Feb 27, 2002 this sequence version replaced gi:16799024.
  ----- Genome Center
  Center: University of Washington Genome Center
  Center Code: UWGC
  Web site: http://www.genome.washington.edu
  Contact: uwgchgs@u.washington.edu
  Drafting Center: SC
  ----- Project Information
  Center project name: chr-1
  Center clone name: RP11-145A3 (sc0642)
  ----- Summary Statistics
  Sequencing vector: plasmid; 44% of reads
  Sequencing method: plasmid; L08752; 56% of reads
  Chemistry: Dye-terminator ET; 68% of reads
  Chemistry: Dye-terminator Big Dye; 32% of reads
  Assembly program: Phrap; version 0.990319
  Consensus quality: 149371 bases at least Q40
  Consensus quality: 149548 bases at least Q40
  Insert size: 40578; sum-of-contigs
  Quality coverage: 28.0x in Q20 bases; sum-of-contigs

```

Overlapping sequences:
 5': Mapping in progress
 3': RP11-92N17 AL513528

Sequence Quality Assessment:
 This entry has been annotated with sequence quality

ATCATGCCAGAGATGCGCTGTGCTTGAAGCAAACTGTGTATCCGCAATCGAAAC

Query Match 9.8%; Score 251.2; DB 9; Length 149548;

Best Local Similarity 80.2%; Pred No. 2.3e-61;

Matches 295; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

```
QY 1778 CTTGGACAGTCTAGGTCCTCAATGTCCCCCATTTGAGAGACAAACAGCCCAAGCTCTTTT 1837
DB 65460 CTGAGGTAGTCCCTGTAGGTAGTGTCTTTTGAAGAGATTAACAATCTTATTTCTT 65401
QY 1838 CTTTTTTTGAAGAGAGTCTTGCCCTGTGTCATGCTGAGTGAATGSCAGATCT 1897
DB 65400 TTTTTTTTGAAGATGAGTCTCTCTCTGTGTGCTCCAGGCTGAGTGAATGSCGCAATCT 65341
QY 1898 CAGCTCAGTCAACCTCCATCTCTGTGATTCAACAAATTCTCTGCTCAGCTTCAAGAA 1957
DB 65340 CGGCTCAGTCAACCTCTGCTCTGTGATTCAAGCAATTCCTGCTCAGCTTCCGAG 65281
QY 1958 TAGCTGGGATTAAGGCGTACACCAAGCTGCTGCTAATTTTTTTTGTATTTTGTAGA 2017
DB 65280 TAGCTGGGATTAAGGCGTACACCAAGCTGCTGCTAATTTTTTTTGTATTTTGTAGA 65221
QY 2018 CATGGGTTTACACATTTGSCCAGGCTGTGCTCAACTCTGACCTCAGTGAATCCACC 2077
DB 65220 GACAGGGTTTCAACATTTGSCCAGGCTGTGCTCAAACTCTGACCTCAGTGAATCCACC 65161
QY 2078 CACCTGGCTCCCAAGTGTGGATTACAGGTGTGAGCCAGCAGCAGCTTAGCTC 2137
DB 65160 CACCTGGCTCCCAAGTGTGGATTACAGGTGTGAGCCAGCAGCAGCTTAGGAG 65101
QY 2138 TCAGATCT 2145
DB 65100 TACCACT 65093
```

Search completed: March 30, 2003, 06:24:26
Job time : 9125.98 secs